[TILTP1] TILASTOTIETEEN JOHDANTOKURSSI, Syksy 2011
http://www.uta.fi/~strale/tiltp1/index.html

HARJOITUS **5** viikko **41**
Joitain ratkaisuja

1. $H_0 : \pi = 10$, $H_1 : \pi \neq 10$. Kun H_0 on tosi, niin

$$Z = \frac{p - \pi_0}{\sqrt{\pi_0(100-\pi_0)/n}} \overset{likimain}{\sim} N(0,1), \text{ joten } z = \frac{9-10}{\sqrt{10(100-10)/200}} = -0.47.$$

Normaalijakauman taulukosta liitteestä 2 saadaan $z_{0.05} = 1.6449 > |-0,47|$, joten H_0 hyväksytään 10 %:n ristitasolla tarkasteltuna. Pienin riskitaso, jolla nollahypoteesisi voidaan hylätä (p-arvo), on P(Z > 0,47 tai Z < -0,47) > 0.10.

2. $H_0 : \mu = 1000$, $H_1 : \mu \neq 1000$. Kun H_0 on tosi, niin $t = \dfrac{\overline{X} - 1000}{s/\sqrt{n}} \sim t(n-1)$.

Aineistosta laskettu $t = \dfrac{1002-1000}{3.4/\sqrt{20}} = 2.63$. Onko saatu t-arvo harvinaisten vai tavanomaisten arvojen joukkoon kuuluva? Verrataan laskettua t-arvo taulukkoarvoihin. Nyt α = 0,01, α/2 = 0,005 ja $t_{0.005;19} = 2{,}861$. Koska 2,63 < 2,861 , niin H_0 hyväksytään 1% riskitasolla tarkasteltuna. Jos riskitaso olisi 2%, niin nollahypoteesi tulisi hylättyä. Siis 0,01 < p < 0,02 (kaksisuuntaisessa testissä).

3. Tutkitaan erotuksia. Erotuksen keskiarvo -2.3 ja keskihajonta 3.4, n = 10.

$H_0 : \mu = 0$, $H_1 : \mu < 0$ (eli B parempi). Kun H_0 on tosi, niin $t = \dfrac{\overline{X} - 0}{s/\sqrt{n}} \sim t(n-1)$.

$t = \dfrac{-2.3-0}{3.4/\sqrt{10}} = -2.14$. Onko saatu t-arvo harvinaisten vai tavanomaisten arvojen joukkoon kuuluva? Verrataan laskettua t-arvo taulukkoarvoihin. Koska -2.14 < -1.833 = $-t_{0.05;9}$, niin H_0 hylätään ja hyväksytään H_1 (siis B parempi) 5 %:n riskitasolla tarkasteltuna. Koska -2.14 >- 2.262 = $-t_{0.025;9}$, niin H_0 hyväksytään (ei eroja) 2.5 %:n riskitasolla tarkasteltuna. Siis p-arvo eli pienin riskitaso, jolla H_0 voidaan hylätä, on 0.025 < p < 0.05 (yksisuuntaisessa testissä).

4.
H_0 : Miesten ja naisten keskimääräinen päivässä saama kolesterolimäärä on sama
H_1 : eivät samoja

Testisuureen arvo on 4.665 (katsotaan 1. riviltä, koska voidaan olettaa populaatioiden varianssit yhtä suuriksi, F-testiin liittyvä p = 0.083) ja H_0 hylätään, koska pienin riskitaso, jolla H_0 voidaan hylätä, on < 0,001 (kolmella desimaalilla 0,000). Päätellään, että keskimääräiset kolesterolimäärät eivät samoja.

5.
H_0 : ei riippuvuutta
H_1 : on riippuvuutta.

Koska p = 0.0142 eli pienin riskitaso, jolla H_0 voidaan hylätä, niin H_0 hylätään tätä suuremmilla riskitasoilla, mutta hyväksytään pienemmillä. Esim. 1 % riskitasolla hyväksytään, mutta 1,5 % riskitasolla hylätään.

6.
H_0 : ei riippuvuutta
H_1 : on riippuvuutta.

Koska p = 0,344>0,05, niin nollahypoteesi hyväksytään eli ei riippuvuutta.

7. Tutkitaan lineaarisen riippuvuuden olemassaoloa. Testattu hypoteesi H_0 : Muuttujien välinen korrelaatiokerroin populaatiossa nolla , voidaan merkitä $\rho = 0$, vaihtoehtoinen hypoteesi $H_1 : \rho \neq 0$, pienin riskitaso, jolla H_0voidaan hylätä on tuloksista kolmen desimaalin tarkkuudella 0,000; voidaan sanoa siis, että $p < 0,001$. H_0 hylätään. Päätellään, että populaatiossa muuttujien välillä lineaarista riippuvuutta.